DOCUMENT RESUME

ED 166 647 CS 004 598

AUTHOR Greenlaw, M. Jean; Moore, David
TITLE An Analysis of Current Research on the Metropolitan Readiness Test and First-Grade Reading Achievement.
PUB DATE 78
NOTE 16p.; Paper presented at the Annual Meeting of the National Reading Conference (28th, St. Petersburg Beach, Florida, November 30-December 2, 1978) ; Not available in hard copy due to marginal legibility

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS *Beginning Reading; Evaluation Criteria; Individualized Instruction; Literature Reviews; *Predictive Validity; Primary Education; *Reading Achievement; Reading Diagnosis; Reading Readiness; *Reading Readiness Tests; *Reading Research; Research Criteria; *Test Validity
IDENTIFIERS *Metropolitan Readiness Test

ABSTRACT

A review of literature by reading authorities identified seven criteria necessary for meaningful research about reading readiness tests. These criteria were then used in analysis of 15 research studies dealing with the "Metropolitan Readiness Test" and first grade reading achievement. Five major findings resulted from this analysis: first, MRT scores provide few direct implications about the instructional programs that individual students should receive (contrasting with the consensus of reading experts that reading readiness tests and research should help to determine how children are ready to begin reading instead of if they are ready to learn to read). Second, the time when the MRT is administered does not substantially affect its predictive validity. Third, research reports ignore aptitude-treatment interactions affecting MRT's validity, making it difficult to determine MRT's usefulness in conjunction with different instructional circumstances. Fourth, the statistical power computed in the studies (indicating MRT's ability to predict or detect significant relationships) is extremely high. Fifth, the average correlation of .59 between the MRT and first grade reading attainment is large enough to warrant predictions of group reading performance but not of individual reading performances. (A bibliography lists the resources used to establish the criteria and to collect the research data.) (RL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166647

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

BEST COPY AVAILABLE

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. Jean Greenlaw
David Moore

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

An Analysis of Current Research on the Metropolitan Readiness Test and First-Grade Reading Achievement

Dr. M. Jean Greenlaw
College of Education
North Texas State University
Denton, TX 76203

Mr. David Moore
309 Aderhold
University of Georgia
Athens, GA 30602

Even a brief review of the research literature on beginning reading readiness confirms Underwood's dictum that "the history of science is the history of relentless analysis" (1957). Demographic, intellectual, perceptual, linguistic, attitudinal, and instructional factors have been relentlessly analyzed individually and in combination to determine their effect on beginning reading success. As of 1972, aspects of some of these factors have been combined into 29 published tests which purport to measure students' readiness for beginning reading instruction (Buros, 1972). Of these 29, the

CS004598

Metropolitan Readiness Test (MRT) (Hildreth, 1964) is used most often in our public schools (Maitland, 1974). The type and pace of beginning reading instruction that numerous children receive is regularly decided, in part at least, by their performance on the MRT. A careful look at current research evaluating the use of the MRT as a pre-reading assessment instrument is therefore warranted.

To provide focus while examining the research articles, specific criteria were set. These criteria were selected from those suggested in the literature by various reading authorities. Several articles were consulted (Baker, 1968; Barrett, 1970; Dykstra, 1967; Farr and Anastasiow, 1969; Hillerich, 1978; MacGinitie, 1969; Rude, 1978; and Wanat, 1976) and the criteria they suggest for research on readiness tests to be meaningful were synthesized. From this synthesis, the following questions were asked:

1. Is the effectiveness of differentiated instruction based on MRT results given?

By far, the most common, urgent recommendation among the reading authorities cited here was to conduct studies wherein students are individually diagnosed, instruction and the criteria of achievement adapted, and the effectiveness of the instructional intervention reported. The predominant call was to determine how children are ready to begin reading instruction -- not if they are ready. If studies were to follow this recommendation, many of the concerns about the MRT's predictive validity would no longer be necessary. Instead, the MRT would be used mostly as a screening device to make tentative decisions about students. After reporting the results of the initial screening, research studies would further investigate the

diagnostic follow-up, the subsequent adaptation of instruction and then later reading achievement. This would all be an attempt to derive instructional implications from readiness testing.

2. Is the time of the MRT administration stated?

Readiness tests are generally administered either at the end of kindergarten or at the beginning of first grade. If researchers choose to study the predictive validity of the MRT, reporting the time when it was administered becomes important. Otherwise, generalizations cannot be safely made.

3. Are the sample subject's characteristics adequately described?

Previous school experience, SES, race, sex, IQ and linguistic performance, among other variables, have differing effects on children's ability to learn to read. It behooves researchers and research journals to provide as much of this information as possible so others can observe the MRT's validity with differing groups of children.

4. Are the instructional methods and materials adequately described?

Certain variables may enhance reading success in one instructional setting but not in another. For instance, ability to blend isolated sounds through auditory and speech channels seems to be more necessary for success in a synthetic phonic program than in a whole-word, analytic approach. In order to evaluate the validity of the MRT with various groups, the treatment they received, along with their other characteristics, must be known.

5. Is the criterion of reading attainment given?

Since reading is such a multi-faceted phenomena, it is necessary to report the criterion measure used to describe it. Otherwise, evaluating its ability to predict other criterion becomes guesswork.

6. Is the MRT's ability to predict an individual student's reading achievement provided?

For accurate prediction of individual's achievement, the correlation between the predictor and the criterion must be quite high. Other statistical information on the prognosis of individuals may be determined by using such techniques as expectancy tables or discriminant analysis.

7. Does the study have adequate power to detect possible relationships?

Statistical power analysis shows whether sampling error regarding beta risk was made. Beta, or Type II, error means one has falsely accepted the null hypothesis. That is, a researcher finds no statistically significant relationship between variables when one, in fact, does exist. Having a small sample, especially, raises the beta risk and, obversely, reduces power.

Summary.

Fifteen original research studies of the MRT published since 1969 were located and examined according to the seven criteria listed above. It is important to note that only studies dealing with the MRT and first-grade reading achievement were chosen for this review. For clarity, the findings are displayed in table form unlike traditional reviews such as Livo's (1972) which describe the highlights of each and every study in running text.

- - - - - - - - -

Insert Table 1

- - - - - - - - -

TABLE 1

1	2	3	4	5	6	7	8		9
Study	Screening	Time of Admin.	N	Subject's Description	Instru. Prog.	Criterion Measure	r	R 1.23	Power
Bolig & Fletcher	No	Post K	625	Location, Sex	No	SAT	.72		.995
						SAT		.73[1]	.995
						1st Gr. Teacher Rdg.	.61		.995
						1st Gr. Teacher Rdg.		.61[2]	.995
Brown	No	Pre 1st	30	Race, Sex, SES Pre-Sch. experience	No	SORT	.46		
Buttram et. al.	No	Post K	97	Location, age, Sex, # siblings, education of parents	No	"1st Grade Reading"	.67		.995
Coffman, et. al.	No	Post K	156	Sex	No	SAT w/boys			
						Word rdg.	.62		
						Para. meaning	.63		
						Vocab.	.59		
						Word study	.68		
						SAT w/girls			
						Word rdg.	.59		
						Para. meaning	.58		
						Vocab.	.56		
						Word study	.59		
Coffman, et. al.	No	Post K	109	Sex	No	SAT			
						Word rdg.	.50		
						Para. meaning	.54		
						Vocab.	.56		
						Word Study	.52		
Engin	No	Post K	100	SES, IQ, age, race	No	G-M: Vocab.		.63[4]	.995
						G-M: Comp.		.58[5]	.995
Goodman & Wiederholt	No	—	49	SES, IQ, sex, Race, pre-school experience	No	SAT: Para. Mng.	.64		.98
						SAT: Vocab.		.71[6]	.995

TABLE 1 (cont.)

1	2	3	4	5	6	7	8		9
Study	Screening	Time of Admin.	N	Subject's Description	Instru. Prog.	Criterion Measure	r	R 1.23	Power
Harckham	No	K	553	Location	No	MAT	.63		
Huberty & Swan	No	Pre-1st	58	.SES, IQ, race, Sex, no pre-sch. experience	No	SAT: wd. rdg.		.75[7]	.995
						SAT: Para. mng.	.71		.995
						SAT: Vocab.		.71[6]	.995
						SAT: Wd. St. SK.		.73[7]	.995
Lewis	No	Pre-1st	40	Location	No	Lee-Clark	.67		.95
						1st Grade Teacher Rdg.	.61		.95
Randal et. al.	No	Post K	62	SES, age	No	SAT: total raw score of rdg. subtests	.34		.41
								.34[8]	.41
Rubin	No	Pre K	910	SES, IQ, sex, location, lang. dev.	No	WRAT	.52		.995
		Pre 1st					.54		.995
Stevenson et. al.	No	Pre 1st	144	SES, boys, parent edu., race	No	WRAT	.44		.995
			151	SES, girls, parent edu., race			.48		.995
Telgedy	No	Post K	56	SES, sex, age	No	WRAT	.70		.995
						GORT: Wd. Recog.	.58		.93
						GORT: Comp.	.67		.95
William	No	Post K	133	SES, race, sex	No	CAT		.65[9]	.995

Column two shows that the criteria most strongly advocated by the reading authorities cited earlier is totally lacking in these studies. None of the original research reviewed here measured the value of the MRT as part of a diagnostic battery for making instructional decisions. In no instance was the MRT used for screening and then followed by diagnostic placement and differentiated instruction of individuals. Because of this, any decisions about individual instruction based on MRT scores would result from teacher intuition only and not from current research evidence.

As column three portrays, every study noted when the MRT was administered. The prevailing practice was to give it at the end of students' kindergarten year. The Rubin study (1974), which compared pre-K with pre-first-grade testing, concluded that the MRT provided basically equivalent predictive results regardless of when it was administered. In order to investigate this further, the correlation coefficients of studies giving the MRT at post-K (spring of the year) were compared with those given pre-first (fall of the year). Using Fisher's Z' transformation to compile the results, the average predictive value of the post-K studies was found to be .54. The average for the pre-first studies was .59. This slight difference tends to confirm Rubin's conclusion that the time of readiness testing has little effect on later prediction of group achievement.

The sample subjects' characteristics which were reported are listed in column 5. As can be seen, this aspect of report writing was given uneven treatment. While Huberty and Swan (1974) extensively reported subjects' characteristics and even controlled pre-school experience to see its effect on the predictive validity of the MRT, Lewis (1974)

simply reported the geographic location of his subjects. All the studies provided some description, but this important aspect seems to have been mostly given cursory attention. Because of this, determining overall aptitude interactions with the MRT becomes difficult.

As column 6 shows, no study provided information about the reading instruction subjects received. One is generally given the correlation between predictor and criterion scores, but left wondering what treatment the subjects received to influence that criterion measure. This aspect of measuring aptitude-treatment interactions has been given no attention by these studies.

Column 7 shows that the criteria of reading achievement had wide variation among the studies. The Buttram (1976) article gave no indication whatsoever of the specific reading measure used. Even though the Stanford Achievement Test was found in three studies, different scoring combinations were used for all three. Bolig and Fletcher (1973) used total SAT scores, Huberty and Swan (1974) worked with individual SAT subtest scores and Randel, et. al. computed a (1977) total raw score for the reading subtests.

As column 8 shows, single order correlation coefficients between the MRT and first grade reading achievement range from .34 to .72. Of the 29 coefficients reported in these studies, the median value was .59 and, interestingly, the mean value computed through Fisher's Z' transformations was also .59. This means that for the 3,122 students involved in the studies compiled here, approximately 35% of their variability in first grade reading achievement can be determined by knowing their MRT scores. While this is a rather substantial figure and indicates that the MRT is appropriate for group evaluation

and prediction, it is still not high enough to allow individual prediction of reading success. There is still much room for individual variation. In addition, only one study, Stevenson (1976), followed up individual students' prognosis to report the "hits", "misses" and "false alarms" made in an attempt to specify the predictive validity of the MRT.

The multiple correlation coefficients range from .75 to zero (non-significant findings). The median value here is .65 and the mean is .62. These findings, based on 11 multiple regression equations reported in the studies; indicate a very slight overall increase from .59. The validity of the MRT seems to be little changed when other variables are entered into the prediction equation. This attests to the relative strength of the MRT to stand by itself in predicting group achievement.

Column 8 lists the power values for each study obtained from Cohen's tables in his Statistical Power Analysis for the Behavioral Sciences (1969). As mentioned earlier, power is an estimation of one's chances of avoiding a Type II error. It especially reflects the sample size of a study, although the effect size and alpha risk also determine its extent. A study should have a minimum power value of .80 to avoid beta error. As shown, all but one study exceeds the .80 minimum, and the majority exceed .995.

Conclusions.

The findings of this review point out several interesting aspects of the current research evidence dealing with the MRT. First of all, few direct implications for the instructional program individuals should receive can be based on their MRT scores.

At least in the studies reported here, no support for such decisions was found. Second, it appears that the time of administration does not substantially affect the predictive validity of this instrument. Third, these research reports have largely ignored looking at aptitude-treatment interactions as they affect the validity of the MRT. Comparisons of the MRT's usefulness in different instructional circumstances are difficult to make since those circumstances are rarely described well. Fourth, the statistical power computed for these studies was extremely high in the great majority of cases. Finally, the average correlation of .59 between the MRT and first-grade reading attainment is large enough to warrant predictions of group reading performance and to make tentative decisions about individual placement. However, even the addition of other predictor variables does not raise the correlational value high enough to allow predictions of individual's performanc

I. Bibliography

Baker, E. L. "Methodological Considerations for Future Readiness Research." Paper presented at annual meeting of the International Reading Association, 1968 (ED 026 208).

Barrett, Thomas. "Predicting Reading Achievement through Readiness Tests." Measurement and Evaluation of Reading, Ed. Roger Farr. New York: Harcourt, Brace and World, Inc., 1970.

Buros, Oscar. The Seventh Mental Measurements Yearbook. Highland Park, New Jersey: Gryphon Press, 1972.

Cohen, Jacob. Statistical Power Analysis for the Behavioral Sciences. New York: Academic Press, 1969.

Dykstra, Robert. "The Use of Reading Readiness Tests for Prediction and Diagnosis: A Critique." The Evaluation of Children's Reading Achievement, Ed. Thomas Barrett. Newark, Delaware: International Reading Association, 1967.

Farr, Roger, and Anastasiow, Nicholas. Tests of Reading Readiness and Achievement: A Review and Evaluation. Newark, Delaware: International Reading Association, 1969.

Hildreth, G., Griffiths, N. and McGauvran, M. Metropolitan Readiness Tests. New York: Harcourt, Brace, Jovanovich, 1964.

Hillerich, Robert L. "A Diagnostic Approach to Early Identification of Language Skills." The Reading Teacher, Vol. 31 (January, 1978), pp. 357-364.

Livo, Norma J. "Reading Readiness: Research in Review" ERIC Clearinghouse on Reading, 1972 (ED 059 854).

MacGinitie, Walter H. "Evaluating Readiness for Learning to Read: A Critical Review and Evaluation of Research." Reading Research Quarterly, Vol. 4 (Spring, 1969), pp. 396-410.

Maitland, Suzanne, Nadeau, J. B. E. and Nadeau, Gretchen. "Early School Screening Practices." Journal of Learning Disabilities, Vol. 7 (December, 1974), pp. 55-59.

Rude, Robert T. "Readiness Tests: Implications for Early Childhood Education." The Reading Teacher, Vol. 26 (March, 1973), pp. 572-580.

Underwood, B. J. Psychological Research. New York: Appleton-Century-Crofts, 1957.

Wanat, Stanley J. "Reading Readiness." Visible Language, Vol. 10 (Spring, 1976), pp. 101-127.

II. Bibliography: Original Research Studies Evaluated

Bolig, John, and, and Fletcher, Gerald. "The MRT vs. Ratings of Kindergarten Teachers as Predictors of Success in First Grade." Educational Leadership, Vol. 30 (April, 1973), pp. 637-639.

Brown, Nina W. "Non-Cognitive Characteristics in the Prediction of Reading Readiness." Educational and Psychological Measurement, Vol. 36 (Summer 1976), pp. 537-542.

Buttram, Joan, Covert, Robert and Hayes, Marjorie. "Prediction of School Readiness and Early Grade Achievement by Classroom Teachers." Educational and Psychological Measurement, Vol. 36 (Summer, 1976), pp. 543-546.

Coffman, Alice, et. al. "Prediction of Achievement in the First Primary Year: Study Number One." 1969 (ED 044 180).

Coffman, Alice, et. al. "Primary Mental Abilities and Metropolitan Readiness Tests as Predictors of Achievement in the First Primary Year." 1970 (ED 043 683).

Engin, Ann. "The Relative Importance of the Subtests of the Metropolitan Readiness Test in the Prediction of First Grade Reading and Arithmetic Achievement Criteria." Journal of Psychology, Vol. 88 (November, 1974), pp. 289-298.

Goodman, Libley and Wiederholt, J. Lee. "Predicting Reading Achievement in Disadvantaged Children." Psychology in the Schools, Vol. 10 (April 1973), pp. 181-184.

Harckham, L. D., et. al. "Multiple Prediction of Reading Achievement in Grades One through Four Using Kindergarten Measures." Paper presented at the annual meeting of the American Educational Research Association, 1971 (ED 049 311).

Huberty, Carl, and Swan, William. "Preschool Classroom Experience and First-Grade Achievement." The Journal of Educational Research, Vol. 67 (March, 1974), pp. 311-316.

Lewis, John. "A Study of the Validity of the Metropolitan Readiness Test." Educational and Psychological Measurement, Vol. 34 (Summer, 1974), pp. 415-416.

Randel, Mildred, Fry, Maurine and Ralls, Elizabeth. "Two Readiness Measures as Predictors of First- and Third-Grade Reading Achievement." Psychology in the Schools, Vol. 14 (January, 1977), pp. 37-40.

Rubin, Rosalyn. "Preschool Application of the Metropolitan Readiness Tests: Validity, Reliability and Preschool Norms." Educational and Psychological Measurement, Vol. 34 (Summer, 1974), pp. 417-422.

Stevenson, Harold, et. al. "Longitudinal Study of Individual Differences in Cognitive Development and Scholastic Achievement." Journal of Educational Psychology, Vol. 68 (August, 1976), pp. 377-400.

Telgedy, Gabriel. "The Effectiveness of Four Readiness Tests as Predictors of First-Grade Academic Achievement." Psychology in the Schools, Vol 12 (January, 1975), pp. 4-11.

Williams, Jean. "The Relationship of Self-Concept and Reading Achievement in First-Grade Children." The Journal of Educational Research, Vol. 66 (April, 1973), pp. 378-380.